

SEQUENCE LISTING

(i) APPLICANT:

- (A) NAME: Kai Krohn et al.
(B) STREET: Iltarusko, Salmentaantie 751
(C) CITY: 36450 Salmentaka
(E) COUNTRY: Finland
(F) POSTAL CODE (ZIP): none

(ii) TITLE OF INVENTION: Novel Gene

(iii) NUMBER OF SEQUENCES: ~~20~~

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2036 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE/:

- (A) NAME/KEY: CDS
(B) LOCATION: 137..1774
(D) OTHER INFORMATION: /product= "AIR-1"

(ix) ~~FEATURE:~~

- ```
(A) NAME/KEY: mat_peptide
(B) LOCATION:137..1771
(D) OTHER INFORMATION:/product= "AIR-1"
```

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AGACCGGGA GACGGGCGGG CGCACAGCCG GCGCGGAGGC CCCACAGCCC CGCCGGGACC 60

CGAGGCCAAG CGAGGGGCTG CCAGTGTCCC GGGACCCACC GCGTCCGCC CAGCCCCGGG 120

TCCCCGCGCC CACCCC ATG GCG ACG GAC GCG GCG CTA CGC CGG CTT CTG 169  
Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu

Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu  
1 5 10

Sub  
B1

Sub B1

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| AGG CTG CAC CGC ACG GAG ATC GCG GTG GCC GTG GAC AGC GCC TTC CCA | 217 |
| Arg Leu His Arg Thr Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro |     |
| 15 20 25                                                        |     |
| CTG CTG CAC GCG CTG GCT GAC CAC GAC GTG GTC CCC GAG GAC AAG TTT | 265 |
| Leu Leu His Ala Leu Ala Asp His Asp Val Val Pro Glu Asp Lys Phe |     |
| 30 35 40                                                        |     |
| CAG GAG ACG CTT CAT CTG AAG GAA AAG GAG GGC TGC CCC CAG GCC TTC | 313 |
| Gln Glu Thr Leu His Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe |     |
| 45 50 55                                                        |     |
| CAC GCC CTC CTG TCC TGG CTG CTG ACC CAG GAC TCC ACA GCC ATC CTG | 361 |
| His Ala Leu Leu Ser Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu |     |
| 60 65 70 75                                                     |     |
| GAC TTC TGG AGG GTG CTG TTC AAG GAC TAC AAC CTG GAG CGC TAT GGC | 409 |
| Asp Phe Trp Arg Val Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly |     |
| 80 85 90                                                        |     |
| CGG CTG CAG CCC ATC CTG GAC AGC TTC CCC AAA GAT GTG GAC CTC AGC | 457 |
| Arg Leu Gln Pro Ile Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser |     |
| 95 100 105                                                      |     |
| CAG CCC CGG AAG GGG AGG AAG CCC CCG GCC GTC CCC AAG GCT TTG GTA | 505 |
| Gln Pro Arg Lys Gly Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val |     |
| 110 115 120                                                     |     |
| CCG CCA CCC AGA CTC CCC ACC AAG AGG AAG GCC TCA GAA GAG GCT CGA | 553 |
| Pro Pro Pro Arg Leu Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg |     |
| 125 130 135                                                     |     |
| GCT GCC GCG CCA GGA GCC CTG ACT CCA AGG GGC ACC GCC AGC CCA GGC | 601 |
| Ala Ala Ala Pro Ala Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly |     |
| 140 145 150 155                                                 |     |
| TCT CAA CTG AAG GCC AAG CCC CCC AAG AAG CCG GAG AGC AGC GCA GAG | 649 |
| Ser Gln Leu Lys Ala Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu |     |
| 160 165 170                                                     |     |
| CAG CAG CGC CTT CCA CTC GGG AAC GGG ATT CAG ACC ATG TCA GCT TCA | 697 |
| Gln Gln Arg Leu Pro Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser |     |
| 175 180 185                                                     |     |
| GTC CAG AGA GCT GTG GCC ATG TCC TCC GGG GAC GTC CCG GGA GCC CGA | 745 |
| Val Gln Arg Ala Val Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg |     |
| 190 195 200                                                     |     |
| GGG GCC GTG GAG GGG ATC CTC ATC CAG CAG GTG TTT GAG TCA GGC GGC | 793 |
| Gly Ala Val Glu Gly Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly |     |
| 205 210 215                                                     |     |
| TCC AAG AAG TGC ATC CAG GTT GGC GGG GAG TTC TAC ACT CCC AGC AAG | 841 |
| Ser Lys Lys Cys Ile Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys |     |
| 220 225 230 235                                                 |     |

23

TTC GAA GAC TCC GGC AGT GGG AAG AAC AAG GCC CGC AGC AGC AGT GCC 889  
Phe Glu Asp Ser Gly Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly  
240 245 250

CCG AAG CCT CTG GTT CGA GCC AAG GGA GCC CAG GGC GCT GCC CCC GGT 937  
Pro Lys Pro Leu Val Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly  
255 260 265

GGA GGT GAG GCT AGG CTG GGC CAG CAG GGC AGC GTT CCC GCC CCT CTG 985  
Gly Gly Glu Ala Arg Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu  
270 275 280

GCC CTC CCC AGT GAC CCC CAG CTC CAC CAG AAG AAT GAG GAC GAG TGT 1033  
Ala Leu Pro Ser Asp Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys  
285 290 295

GCC GTG TGT CGG GAC GGC GGG GAG CTC ATC TGC TGT GAC GGC TGC CCT 1081  
Ala Val Cys Arg Asp Gly Gly Gln Leu Ile Cys Cys Asp Gly Cys Pro  
300 305 310 315

CGG GCC TTC CAC CTG GCC TGC CTG TCC CCT CCG CTC CGG GAG ATC CCC 1129  
Arg Ala Phe His Leu Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro  
320 325 330

AGT GGG ACC TGG AGG TGC TCC AGC TGC CTG CAG GCA ACA GTC CAG GAG 1177  
Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu  
335 340 345

GTG CAG CCC CGG GCA GAG GAG CCC CGG CCC CAG GAG CCA CCC GTG GAG 1225  
Val Gln Pro Arg Ala Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu  
350 355 360

ACC CCG CTC CCC CCG GGG CTT AGG TCG GCG GGA GAG GAG GTA AGA GGT 1273  
Thr Pro Leu Pro Pro Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly  
365 370 375

CCA COT GGG GAA CCC CTA GCC GGC ATG GAC ACG ACT CTT GTC TAC AAG 1321  
Pro Pro Gly Glu Pro Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys  
380 385 390 395

CAC CTG CCG GCT CCG CCT TCT GCA GCC CCG CTG CCA GGG CTG GAC TCC 1369  
His Leu Pro Ala Pro Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser  
400 405 410

TCG GCC CTG CAC CCC CTA CTG TGT GTG GGT CCT GAG GGT CAG CAG AAC 1417  
Ser Ala Leu His Pro Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn  
415 420 425

CTG GCT CCT GGT GCG CGT TGC GGG GTG TGC GGA GAT GGT ACG GAC GTG 1465  
Leu Ala Pro Gly Ala Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val  
430 435 440

"SECRET" 85930560

Sub  
B1

24

CTG CGG TGT ACT CAC TGC GCC GCT GCC TTC CAC TGG CGC TGC CAC TTC 1513  
 Leu Arg Cys Thr His Cys Ala Ala Ala Phe His Trp Arg Cys His Phe  
 445 450 455

CCA GCC GGC ACC TCC CGG CCC GGG ACG GGC CTG CGC TGC AGA TCC TGC 1561  
 Pro Ala Gly Thr Ser Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys  
 460 465 470 475

TCA GGA GAC GTG ACC CCA GCC CCT GTG GAG GGG GTG CTG GCC CCC AGC 1609  
 Ser Gly Asp Val Thr Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser  
 480 485 490

CCC GCC CGC CTG GCC CCT GGG CCT GCC AAG GAT GAC ACT GCC AGT CAC 1657  
 Pro Ala Arg Leu Ala Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His  
 495 500 505

GAG CCC GCT CTG CAC AGG GAT GAC CTG GAG TCC CTT CTG AGC GAG CAC 1705  
 Glu Pro Ala Leu His Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His  
 510 515 520

ACC TTC GAT GGC ATC CTG CAG TGG GCC ATC CAG AGC ATG GCC CGT CCG 1753  
 Thr Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro  
 525 530 535

GCG GCC CCC TTC CCC TCC TGA CCCCAGATGG CCGGGACATG CAGCTCTGAT 1804  
 Ala Ala Pro Phe Pro Ser \*  
 540 545

GAGAGAGTGC TGAGAAGGAC ACCTCCTTCC TCAGTCCTGG AAGCCGGCCG GCTGGGATCA 1864

AGAAGGGGAC AGCGCCACCT CTGTCTAGTG CTCGGCTGTA AACAGCTCTG TGTCTCTGGG 1924

GACACCAGCC ATCATGTGCC TGGAAATTAA ACCCTGCCCC ACTTCTCTAC TCTGGAAGTC 1984

CCCGGGAGCC TCTCCTTGCC TGGTGACCTA CTAATAATAT AAAAATTAGC TG 2036

## (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 545 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg Thr  
 1 5 10 15

Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu Leu His Ala Leu  
 20 25 30

Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu His  
 35 40 45

09508668-11000

Sub  
B1



25

Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu/Ser  
50 55 60

Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp Phe Trp Arg Val  
65 70 75 80

Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg Leu Gln Pro Ile  
85 90 95

Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln Pro Arg Lys Gly  
100 105 110

Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro Pro Pro Arg Leu  
115 120 125

Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala Ala Ala Pro Ala  
130 135 140

Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser Gln Leu Lys Ala  
145 150 155 160

Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln Gln Arg Leu Pro  
165 170 175

Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val Gln Arg Ala Val  
180 185 190

Ala Met Ser Ser Gly Asp / Val Pro Gly Ala Arg Gly Ala Val Glu Gly  
195 200 205

Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser Lys Lys Cys Ile  
210 215 220

Gln Val Gly Gly ~~Glu~~ Phe Tyr Thr Pro Ser Lys Phe Glu Asp Ser Gly  
225 230 235 240

Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro Lys Pro Leu Val  
245 250 255

Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly Gly Glu Ala Arg  
260 265 270

Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala Leu Pro Ser Asp  
275 280 285

Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp  
290 295 300

Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu  
305 310 315 320

Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg  
325 330 335

[illegible]

Sub  
B1

26

Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala  
 340 345 350  
 Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro  
 355 360 365  
 Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro  
 370 375 380  
 Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro  
 385 390 395 400  
 Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro  
 405 410 415  
 Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala  
 420 425 430  
 Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His  
 435 440 445  
 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Ala Gly Thr Ser  
 450 455 460  
 Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr  
 465 470 475 480  
 Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala  
 485 490 495  
 Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu Pro Ala Leu His  
 500 505 510  
 Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile  
 515 520 525  
 Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro  
 530 535 540  
 Ser \*  
 545

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1545 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

090883-1000

Sub  
B1

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION:237..1283  
 (D) OTHER INFORMATION:/product= "AIR-2"

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide  
 (B) LOCATION:237..1280  
 (D) OTHER INFORMATION:/product= "AIR-2"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AGAGAAAGTG AGGTCTTCTC AGGCTCTTAA GAGCATGGCG TTTGGTCCAG GCTGTACCCG 60  
 CTGCTCTCAG CTGGGCCCGT GGGTGGCCCG GGCGCCCTG CTATAGCCAG GAGGTCAAGG 120  
 ATCCACTGGG AATGCCATGC TCATCTTTCG TCCCCAGCAT GGTTCCTTAA TGGGGTAGAA 180  
 GCAGGTCGGG AGAGACCTCC CTGGGCCTGG CCOCCTGCC CTGTGAGGAA GGGTTC 236  
 ATG TGG TTG GTG TAC AGT TCC GGG GGC CCT GGA ACG CAG CAG CCT GCA 284  
 Met Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala  
 1 5 10 15  
 AGA AAC CGG GTT TTC TTC CCA ATA GGG ATG GCC CCG GGG GGT GTC TGT 332  
 Arg Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys  
 20 25 30  
 TGG AGA CCA GAT GGA TGG GGA ACA GGT GGT CAG GGC AGA ATT TCA GGC 380  
 Trp Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly  
 35 40 45  
 CCT GGC AGC ATG GGA GCA GGG CAG AGA CTG GGG AGT TCA GGT ACC CAG 428  
 Pro Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln  
 50 55 60  
 AGA TGC TGC TGG GGG AGC TGT TTT GGG AAG GAG GTG GCT CTC AGG AGG 476  
 Arg Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg  
 65 70 75 80  
 GTG CTG CAC CCC AGC CCA GTC TGC ATG GGC GTC TCT TGC CTG TGC CAG 524  
 Val Leu His Pro Ser Pro Val Cys Met Gly Val Ser Cys Leu Cys Gln  
 85 90 95  
 AAG AAT GAG GAC GAG TGT GCC GTG TGT CGG GAC GGC GGG GAG CTC ATC 572  
 Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp Gly Gly Glu Leu Ile  
 100 105 110  
 TGC TGT GAC GGC TGC CCT CGG GCC TTC CAC CTG GCC TGC CTG TCC CCT 620  
 Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu Ala Cys Leu Ser Pro  
 115 120 125

0950666-1000

Sub  
B1

[illegible]

|                                                                 |      |
|-----------------------------------------------------------------|------|
| CCG CTC CGG GAG ATC CCC AGT GGG ACC TGG AGG TGC TCC AGC TGC CTG | 668  |
| Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu |      |
| 130 135 140                                                     |      |
| CAG GCA ACA GTC CAG GAG GTG CAG CCC CGG GCA GAG GAG CCC CGG CCC | 716  |
| Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala Glu Glu Pro Arg Pro |      |
| 145 150 155 160                                                 |      |
| CAG GAG CCA CCC GTG GAG ACC CCG CTC CCC CCG GGG CTT AGG TCG GCG | 764  |
| Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro Gly Leu Arg Ser Ala |      |
| 165 170 175                                                     |      |
| GGA GAG GAG GTA AGA GGT CCA CCT GGG GAA CCC CTA GCC GGC ATG GAC | 812  |
| Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro Leu Ala Gly Met Asp |      |
| 180 185 190                                                     |      |
| ACG ACT CTT GTC TAC AAG CAC CTG CCG GCT CCG CCT TCT GCA GCC CCG | 860  |
| Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro Pro Ser Ala Ala Pro |      |
| 195 200 205                                                     |      |
| CTG CCA GGG CTG GAC TCC TCG GCC CTG CAC CCC CTA CTG TGT GTG GGT | 908  |
| Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro Leu Leu Cys Val Gly |      |
| 210 215 220                                                     |      |
| CCT GAG GGT CAG CAG AAC CTG GCT CCT GGT GCG CGT TGC GGG GTG TGC | 956  |
| Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala Arg Cys Gly Val Cys |      |
| 225 230 235 240                                                 |      |
| GGA GAT GGT ACG GAC GTG CTG CGG TGT ACT CAC TGC GCC GCT GCC TTC | 1004 |
| Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His Cys Ala Ala Ala Phe |      |
| 245 250 255                                                     |      |
| CAC TGG CGC TGC CAC TTC CCA GCC GGC ACC TCC CGG CCC GGG ACG GGC | 1052 |
| His Trp Arg Cys His Phe Pro Ala Gly Thr Ser Arg Pro Gly Thr Gly |      |
| 260 265 270                                                     |      |
| CTG CGC TGC AGA TCC TGC TCA GGA GAC GTG ACC CCA GCC CCT GTG GAG | 1100 |
| Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr Pro Ala Pro Val Glu |      |
| 275 280 285                                                     |      |
| GGG GTG CTG GCC CCC AGC CCC GCC CGC CTG GCC CCT GGG CCT GCC AAG | 1148 |
| Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala Pro Gly Pro Ala Lys |      |
| 290 295 300                                                     |      |
| GAT GAC ACT GCC AGT CAC GAG CCC GCT CTG CAC AGG GAT GAC CTG GAG | 1196 |
| Asp Asp Thr Ala Ser His Glu Pro Ala Leu His Arg Asp Asp Leu Glu |      |
| 305 310 315 320                                                 |      |
| TCC CTT CTG AGC GAG CAC ACC TTC GAT GGC ATC CTG CAG TGG GCC ATC | 1244 |
| Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile Leu Gln Trp Ala Ile |      |
| 325 330 335                                                     |      |
| CAG AGC ATG GCC CGT CCG GCG GCC CCC TTC CCC TCC TGA CCCCAGATGG  | 1293 |
| Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro Ser *               |      |
| 340 345                                                         |      |

CCGGGACATG CAGCTCTGAT GAGAGAGTGC TGAGAAGGAC ACCTCCTTCC TCAGTCCTGG 1353  
 AAGCCGGCCG GCTGGGATCA AGAAGGGGAC AGCGCCACCT CTTGTCAGTG CTCGGGTGTA 1413  
 AACAGCTCTG TGTCTCTGGG GACACCAGCC ATCATGTGCC TGGAAATTAA ACCTGCCCC 1473  
 ACTTCTCTAC TCTGGAAGTC CCCGGGAGCC TCTCCTTGCC TGGTGACCTA CTAAAAATAT 1533  
 AAAAATTAGC TG 1545

## (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 348 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala  
 1 5 10 15  
 Arg Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys  
 20 25 30  
 Trp Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly  
 35 40 45  
 Pro Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln  
 50 55 60  
 Arg Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg  
 65 70 75 80  
 Val Leu His Pro Ser Pro Val Cys Met Gly Val Ser Cys Leu Cys Gln  
 85 90 95  
 Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp Gly Gly Glu Leu Ile  
 100 105 110  
 Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu Ala Cys Leu Ser Pro  
 115 120 125  
 Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu  
 130 135 140  
 Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala Glu Glu Pro Arg Pro  
 145 150 155 160

DEPT 2580560

Sub  
B1

Sub  
B1

(ix) FEATURE:

- (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide  
(B) LOCATION:237..998  
(D) OTHER INFORMATION:/product= "AIR-3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| AGAGAAAGTG AGGTCTTCTC AGGCTCTTAA GAGCATGGCG TTTGGTCCAG GCTGTACCCG  | 60  |
| CTGCTCTCAG CTGGGCCCCGT GGGTGGGCCG GGC GCCCTG CTATAGCCAG GAGGTCAAGG | 120 |
| ATCCACTGGG AATGCCATGC TCATCTTTTCG TCCCCAGCAT GGTTTCTTAA TGGGGTAGAA | 180 |
| GCAGGTCGGG AGAGACCTCC CTGGGCCTGG CCCCCTGCC CTGTGAGGAA GGGTTC       | 236 |
| ATG TGG TTG GTG TAC AGT TCC GGG GCC CCT GGA ACG CAG CAG CCT GCA    | 284 |
| Met Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala    |     |
| 1 5 10 15                                                          |     |
| AGA AAC CGG GTT TTC TTC CCA ATA GGG ATG GCC CCG GGG GGT GTC TGT    | 332 |
| Arg Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys    |     |
| 20 25 30                                                           |     |
| TGG AGA CCA GAT GGA TGG GGA ACA GGT GGT CAG GGC AGA ATT TCA GGC    | 380 |
| Trp Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly    |     |
| 35 40 45                                                           |     |
| CCT GGC AGC ATG GGA GCA GGG CAG AGA CTG GGG AGT TCA GGT ACC CAG    | 428 |
| Pro Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln    |     |
| 50 55 60                                                           |     |
| AGA TGC TGC TGG GGG AGC TGT TTT GGG AAG GAG GTG GCT CTC AGG AGG    | 476 |
| Arg Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg    |     |
| 65 70 75 80                                                        |     |
| GTG CTG CAC CCC AGC CCA GTC TGC ATG GGC GTC TCT TGC CTG TGC CAG    | 524 |
| Val Leu His Pro Ser Pro Val Cys Met Gly Val Ser Cys Leu Cys Gln    |     |
| 85 90 95                                                           |     |
| AAG AAT GAG GAC GAG TGT GCC GTG TGT CGG GAC GGC GGG GAG CTC ATC    | 572 |
| Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp Gly Gly Glu Leu Ile    |     |
| 100 105 110                                                        |     |
| TGC TGT GAC GGC TGC CCT CGG GCC TTC CAC CTG GCC TGC CTG TCC CCT    | 620 |
| Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu Ala Cys Leu Ser Pro    |     |
| 115 120 125                                                        |     |
| CCG CTC CGG GAG ATC CCC AGT GGG ACC TGG AGG TGC TCC AGC TGC CTG    | 668 |
| Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu    |     |
| 130 135 140                                                        |     |
| CAG GCA ACA GTC CAG GAG GTG CAG CCC CGG GCA GAG GAG CCC CGG CCC    | 716 |
| Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala Glu Glu Pro Arg Pro    |     |
| 145 150 155 160                                                    |     |
| CAG GAG CCA CCC GTG GAG ACC CCG CTC CCC CCG GGG CTT AGG TCG GCG    | 764 |
| Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro Gly Leu Arg Ser Ala    |     |
| 165 170 175                                                        |     |

[illegible]

32

GGA GAG GAG CCC CGC TGC CAG GGC TGG ACT CCT CGG CCC TGC ACC CCC 812  
 Gly Glu Glu Pro Arg Cys Gln Gly Trp Thr Pro Arg Pro Cys Thr Pro  
 180 185 190  
 TAC TGT GTG TGG GTC CTG AGG GTC AGC AGA ACC TGG CTC CTG GTG CGC 860  
 Tyr Cys Val Trp Val Leu Arg Val Ser Arg Thr Trp Leu Leu Val Arg  
 195 200 205  
 GTT GCG GGG TGT GCG GAG ATG GTA CGG ACG TGC TGC GGT GTA CTC ACT 908  
 Val Ala Gly Cys Ala Glu Met Val Arg Thr Cys Cys Gly Val Leu Thr  
 210 215 220  
 GCG CCG CTG CCT TCC ACT GGC GCT GCC ACT TCC CAG CCG GCA CCT CCC 956  
 Ala Pro Leu Pro Ser Thr Gly Ala Ala Thr Ser Gln Pro Ala Pro Pro  
 225 230 235 240  
 GGC CCG GGA CGG GCC TGC GCT GCA GAT CCT CCT CAG GAG ACG TGA 1001  
 Gly Pro Gly Arg Ala Cys Ala Ala Asp Pro Ala Gln Glu Thr \*  
 245 250 255  
 CCCCAGCCCC TGTGGAGGGG GTGCTGGCCC CCGCCCCGC CCGCCTGGCC CCTGGGCCTG 1061  
 CCAAGGATGA CACTGCCAGT CACGAGCCCG CTCTGCACAG GGATGACCTG GAGTCCCTTC 1121  
 TGAGCGAGCA CACCTTCGAT GGCATCCTGC AGTGGGCCAT CCAGAGCATG GCCCGTCCGG 1181  
 CGGCCCCCTT CCCCTCCTGA CCCCAGATGG CCGGGACATG CAGCTCTGAT GAGAGAGTGC 1241  
 TGAGAAGGAC ACCTCCTTCC TCACTCCTGG AAGCCGGCCG GCTGGGATCA AGAAGGGGAC 1301  
 AGCGCCACCT CTTGTCACTG CTCGGCTGTA AACAGCTCTG TGTTTCTGGG GACACCAGCC 1361  
 ATCATGTGCC TGGAAATTAA ACCCTGCCCC ACTTCTCTAC TCTGGAAGTC CCCGGGAGCC 1421  
 TCTCCTTGCC TGGTGAOCTA CTAAAAATAT AAAAATTAGC TG 1463

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 254 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala  
 1 5 10 15  
 Arg Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys  
 20 25 30  
 Trp Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly  
 35 40 45

09508658-10300

Sub  
B1

33

Pro Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln  
 50 55 60  
 Arg Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg  
 65 70 75 80  
 Val Leu His Pro Ser Pro Val Cys Met Gly Val Ser Cys Leu Cys Gln  
 85 90 95  
 Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp Gly Gly Glu Leu Ile  
 100 105 110  
 Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu Ala Cys Leu Ser Pro  
 115 120 125  
 Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu  
 130 135 140  
 Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala Glu Glu Pro Arg Pro  
 145 150 155 160  
 Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro Gly Leu Arg Ser Ala  
 165 170 175  
 Gly Glu Glu Pro Arg Cys Gln Gly Trp Thr Pro Arg Pro Cys Thr Pro  
 180 185 190  
 Tyr Cys Val Trp Val Leu Arg Val Ser Arg Thr Trp Leu Leu Val Arg  
 195 200 205  
 Val Ala Gly Cys Ala Glu Met Val Arg Thr Cys Cys Gly Val Leu Thr  
 210 215 220  
 Ala Pro Leu Pro Ser Thr Gly Ala Ala Thr Ser Gln Pro Ala Pro Pro  
 225 230 235 240  
 Gly Pro Gly Arg Ala Cys Ala Ala Asp Pro Ala Gln Glu Thr \*  
 245 250 255

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GATGACACTG CCAATCACGA

00508558.10300

Sub  
B1

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID/NO: 8:

GTTCCCGAGT GGAAGGCGCT GC

22

(2) INFORMATION FOR SEQ ID NO: 9/

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AGGGGACAGG CAGGCCAGGT

20

(2) INFORMATION FOR/SEQ ID NO: 10:

- (i) ~~SEQUENCE/CHARACTERISTICS:~~
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) ~~STRANDEDNESS:~~ single
  - (D) ~~TOPOLOGY:~~ linear

(xi) ~~SEQUENCE~~ DESCRIPTION: SEQ ID NO: 10:

GAGTTCAGGT / ACCCAGAGAT GCTG

24

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTCGCTCAGA AGGGACTCCA

20

**F**

Sub B1

## (2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGATTCAGAC CATGTCAGCT TCA

23

## (2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GTGCTGTTCAGGACTACAA C

21

## (2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TGGATGAGGA TCCCTCCAC G

21

## (2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CCATCCTAAT ACGACTCACT ATAGGGC

27

09508658 110300

Sub  
B1



Sub  
B

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID/NO: 20:

ACGGGCTCCT CAAACACCAC T

21

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

TGGAGATGGG CAGGCCGCAG GGTG

24

(2) INFORMATION FOR ~~SEQ~~ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CAGTCCAGCT GGGCTGAGCA GGTG

24

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GCGGCTCCAA GAAGTGCATC CAGG

24

[illegible]

38

## (2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CTCCACCCTG CAAGGAAGAG GGGC

24

## (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Leu | His | Leu | Lys | Glu | Lys | Glu | Gly | Cys | Pro | Gln | Ala | Phe | His |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     | 15  |     |

## (2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Lys | Asn | Lys | Ala | Arg | Ser | Ser | Ser | Gly | Pro | Lys | Pro | Leu | Val |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     | 15  |     |

## (2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

DEPOSIT 35980660

Sub  
B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

ATGGCGACGG ACGCGGCGCT ACGC

(2) INFORMATION FOR SEQ ID NO: 28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CCTGGATGTA CTTCTTGAG CCGC

(2) INFORMATION FOR SEQ ID NO: 29:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

GAGCCCGAGG GGCCGTGGAG GGGA

(2) INFORMATION FOR SEQ ID NO: 30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GGCTGCACCT CCTGGACTGT TGCC

(2) INFORMATION FOR SEQ ID NO: 31:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

09508658 110300

B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

GATCCTGCTC AGGAGACGTG ACCC

(2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

CACCAGGCAA GGAGAGGCTC CCGG

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CCACCCCATG GCGACGGACG

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

GGAATTCGGA GGGGAAGGGG GCCGCCGGA

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

00508658 110300

B1

